AD 271 606

Reproduced
by the

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA

B O L T   B E R A N E K   A N D   N E W M A N   I N C

C O N S U L T I N G · D E V E L O P M E N T · A P P L I E D   R E S E A R C H

ESD-TDR-62-4

A TEST OF THE 20-BAND AND OCTAVE-BAND METHODS OF COMPUTING THE ARTICULATION INDEX

October 1961

Submitted to:

Operational Applications Laboratory
Electronics Systems Division
L. G. Hanscom Field
Bedford, Massachusetts

ESD-TDR-62-4

A Test of the 20-Band and Octave-Band Methods of Computing the Articulation Index

K. D. Kryter, Gail Flanagan, and Carl Williams
Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge 38, Massachusetts

October 1961

Contract AF 19(604)-4061

Operational Applications Laboratory
Electronics Systems Division
L. G. Hanscom Field
Bedford, Massachusetts

TEST OF THE 20 BAND AND OCTAVE BAND METHODS OF COMPUTING THE ARTICULATION INDEX

Abstract

The calculation of the Articulation Index (AI) as originally formulated requires the preparation of a spectrum level analysis (energy per cycle) of the speech signal coming over a communication system and the noise in which the speech is imbedded. The speech and noise spectra are then divided into 20 narrow bands of frequencies each contributing equally to speech intelligibility. Unfortunately, the making of spectrum level analyses is a relatively laborious process and requires highly specialized laboratory equipment. For this reason it has been proposed that the calculation procedure for AI be modified to permit the use of octave band information as a substitute for spectrum level information (the equipment required to make octave band analyses is commonly available for use in the field).

The studies herein reported were designed to test the accuracy with which AI's calculated by both the octave band and 20 band methods predict the intelligibility of speech presented to listeners in the presence of a variety of types of broad band noises. The results indicate that, for the noises tested, the octave band method for the calculation of AI can be used in place of the more detailed 20 band method without any appreciable loss in the accuracy with which speech intelligibility test scores are predicted.

Reviewed and approved for publication

Herbert Rubenstein

HERBERT RUBENSTEIN, Chief
Information Presentation Division
Operational Applications Laboratory

William V. Hagan

WILLIAM V. HAGAN, Colonel, USAF
Director
Operational Applications Laboratory

A TEST OF THE 20-BAND AND OCTAVE-BAND METHODS OF COMPUTING THE ARTICULATION INDEX

Karl Kryter, Gail Flanagan, and Carl Williams

The so-called Articulation Index (AI), developed by French and Steinberg (1), is considered to be a fairly accurate way of predicting, from purely physical measures, what the intelligibility of speech will be when transmitted over a given system (1, 2, 3, 4). Because the AI technique as originally formulated requires somewhat difficult and detailed physical measures and computations, various attempts have been made to simplify these procedures (2, 5, 6). The general purpose of the present study is to compare the relative accuracy of two methods of calculating this index of speech intelligibility: one is a simplified octave-band method (5, 6), and the other is the older and more basic 20-band method (1, 2).

A large number of speech intelligibility tests have been used for evaluating the accuracy of the Articulation Index in predicting effects of filtering (1, 4, 7, 8) and the effects of masking speech with noise of different band width and location on the frequency scale (9). New studies, however, have attempted to evaluate the accuracy with which the Articulation Index predicts the masking effects of continuous spectrum noises. Pickett and Kryter (6) obtained some data which appeared to show that the AI underestimated the masking effects of a noise which had a spectrum with a positive slope, that is, a spectrum in which the intensity level per cycle increased with increasing frequency.* Egan and

* A check of the test conditions subsequent to the publication of the Pickett and Kryter study revealed a spectrum which differed somewhat from that found in the earlier calibration measurements made of the high frequency noise in question. This, of course, casts doubt on the validity of the results obtained with the "high frequency" noise.

and Thwing (15) in a later study did not find this discrepancy between predicted and obtained results for a noise with a positive sloping spectrum. Tests are included in the present experiment that may help resolve this apparent conflict between these results of the two studies.

PROCEDURE

Six college students served as listeners for the speech intelligibility tests. One talker was used. All 20 Harvard PB word lists (50 words per list) were recorded and four scramblings of each list were available. As the result of participation in other experiments, the listeners had been thoroughly trained on the test material. Two 50-word tests were read at each condition tested with the listeners seated in a soundproofed room.

The tests were presented binaurally over TDH-39 earphones manufactured by the Telephonics Corporation. The average response characteristics of these earphones is shown in Fig. 1. A block diagram of the equipment used is given in Fig. 2.

Four different noise spectra, designated A, B, C, and D were presented electrically to the earphones of the listeners along with the speech signal. Two noise levels, 80 and 105 db re 0.0002 microbar, were used for noise spectra B, C, and D. Noise A was presented only at 80 db. The intensity of the speech signal was systematically varied to obtain several different, but comparable, signal-to-noise ratios for each of the two noise levels. Because of interactions between certain of the noise spectra and the characteristics of the equipment it was necessary to present some of the noises at slightly different absolute levels than others. This range among the different noises around the average levels of 80 and 105 db was only 10 db.

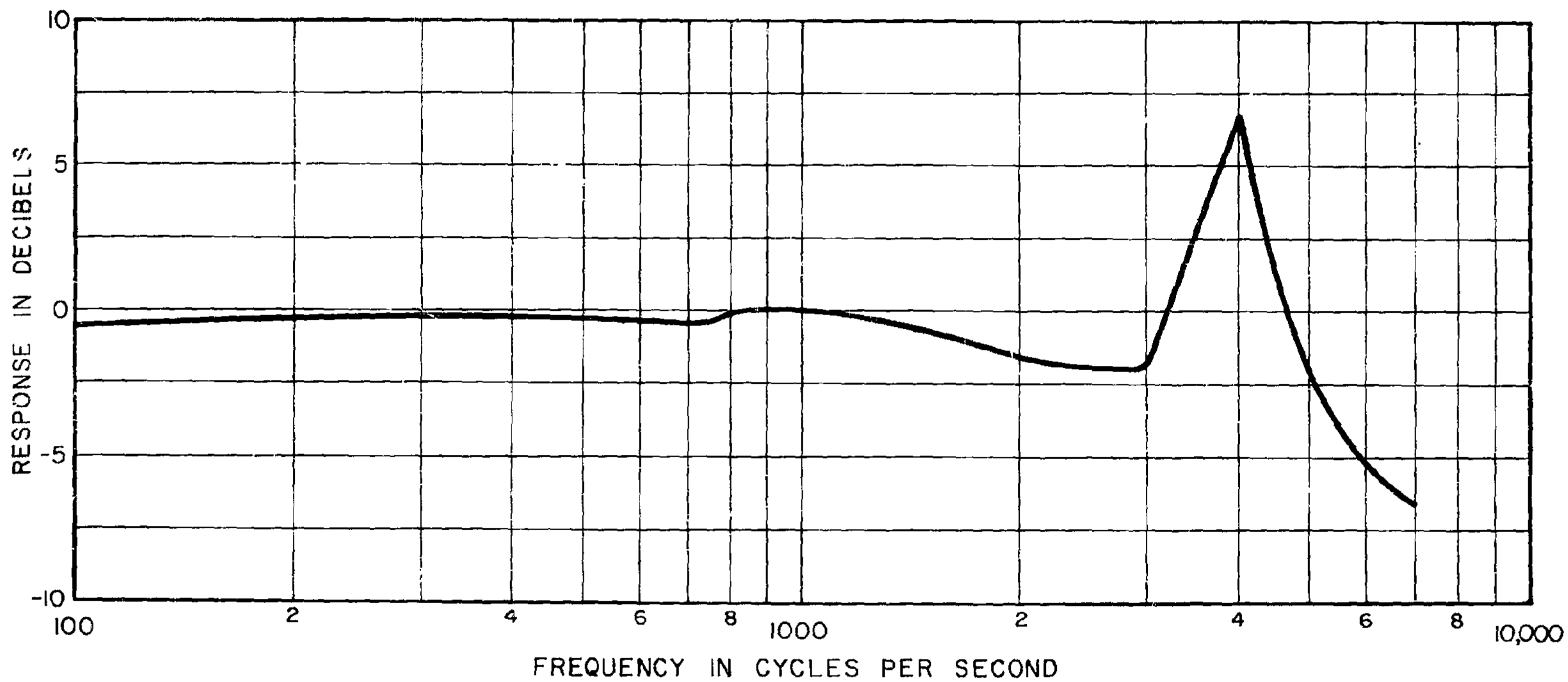

FIG. I AVERAGE RESPONSE OF 12 TDH 39 EARPHONES.
6 CC COUPLER. CONSTANT VOLTAGE INPUT (0.018 VOLT)

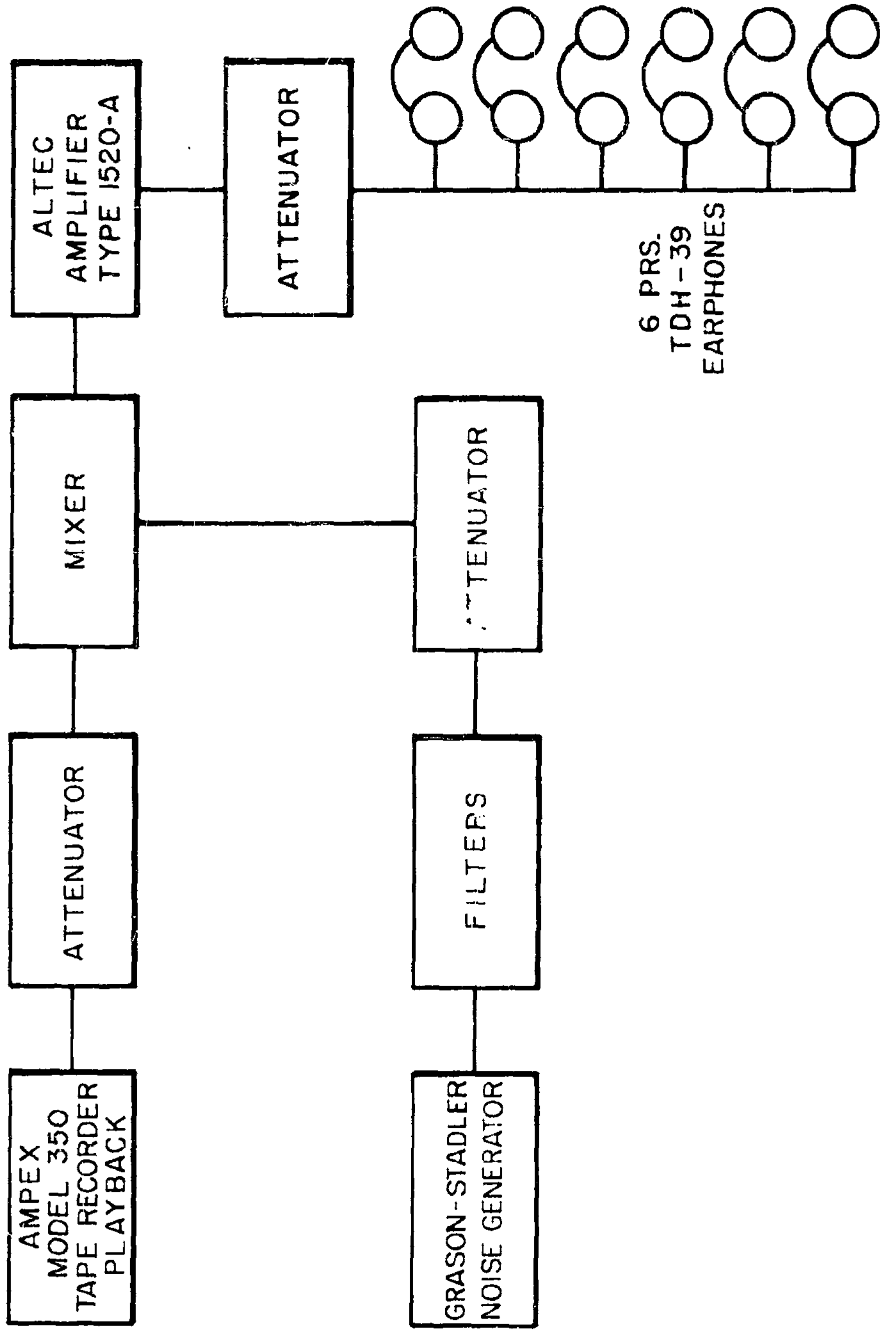

FIG. 2 BLOCK DIAGRAM OF EQUIPMENT USED FOR TESTS

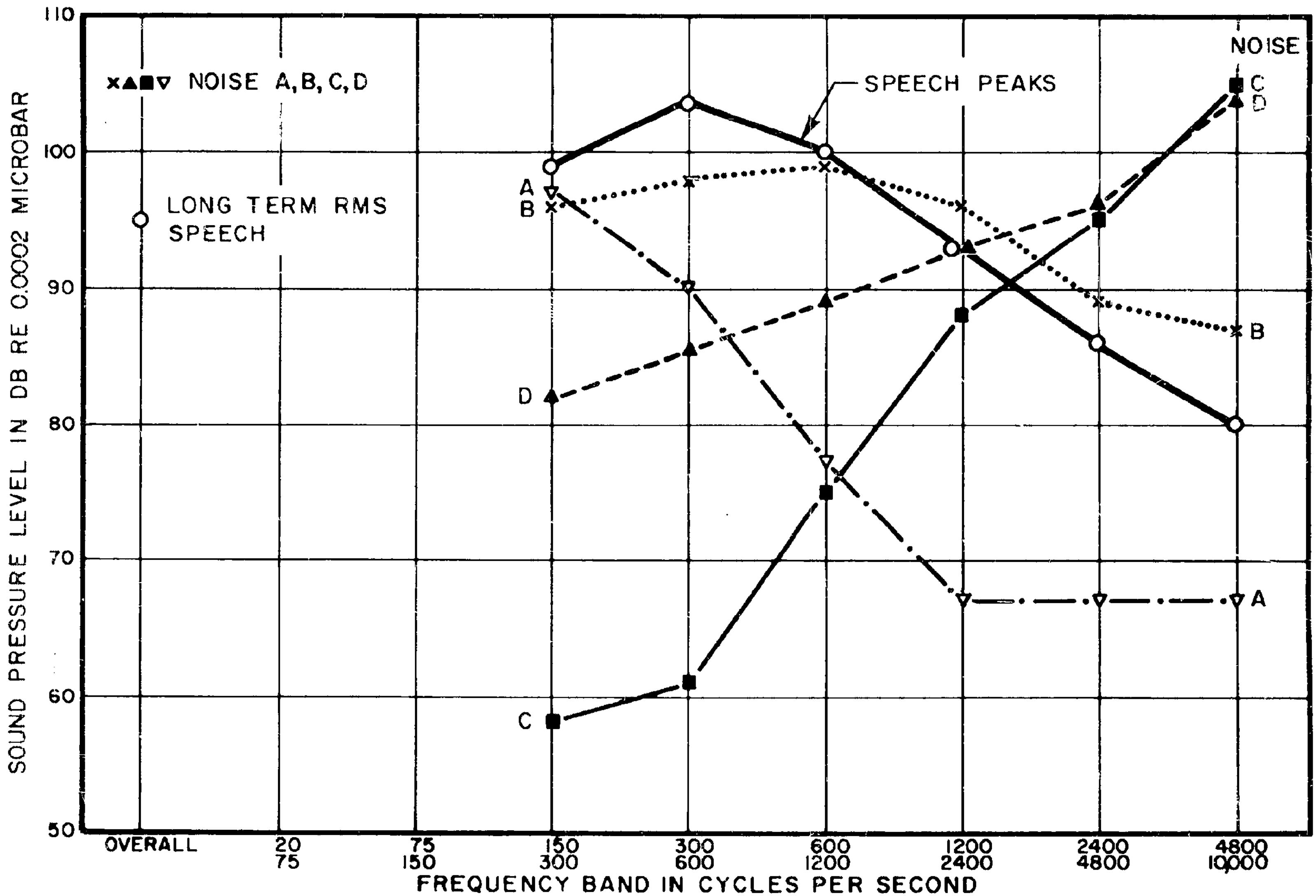

FIG. 3 NOISE AND SPEECH SPECTRA

Examples of the noise and speech spectra used in the tests are shown in Fig. 3. The spectra represent an analysis of the electrical signals applied across an earphone.

Calculation of AI, 20-Band Method

AI's were calculated for the various communication conditions tested according to the 20-band formula suggested by Beranek (2), using the cutoff frequencies given in Table 1. Beranek's formula is

$$AI = \sum_{1}^{20} \left(\frac{S/N}{30.20}\right)$$

where S, the speech peaks, equals in each of the individual 20 bands the long-term root-mean-square (rms) of speech plus 12 db, and N is long-term rms of the noise. This formulation of AI holds that the total fractional value of each of the 20 bands is a linear function of the decibel speech peak-to-noise ratio in that band from a minimum for ratios of and less than 0 db to a maximum for ratios of and greater than 30 db. The maximum value of AI taken over all 20 bands equals 1.0; therefore, by definition the maximum contribution of each of the 20 bands is the fraction .05. Each decibel the speech peak exceeds the noise up to 30 db in each band is accordingly worth .0017 of the total AI (.05/30).

It is often useful when computing AI's by the 20-band method to use the work sheet given in Fig. 4. The speech peak-to-noise ratios for each of the 20 bands can be read directly from spectra plotted on the work sheet. The speech peaks plotted on Fig. 4 are for speech having a long-term rms value 65 db re 0.0002 microbar over all frequencies. This speech level is typical of that used with communication systems under relatively quiet conditions. The overall long-term rms of a speech signal under test can be approximated by subtracting 3 db from the arithmetic average of

Table 1 (after Beranek)

Frequency Bands that Contribute Equally to Speech Intelligibility

Band No.	Lower Frequency	Upper Frequency	Mid-Frequency	Bandwidth
1	200	330	270	130
2	330	430	380	100
3	430	560	490	130
4	560	700	630	140
5	700	840	770	140
6	840	1000	920	160
7	1000	1150	1070	150
8	1150	1310	1230	160
9	1310	1480	1400	170
10	1480	1660	1570	180
11	1660	1830	1740	170
12	1830	2020	1920	190
13	2020	2240	2130	220
14	2240	2500	2370	260
15	2500	2820	2660	320
16	2820	3200	2900	380
17	3200	3650	3400	450
18	3650	4250	3950	600
19	4250	5050	4650	800
20	5050	6100	5600	1050

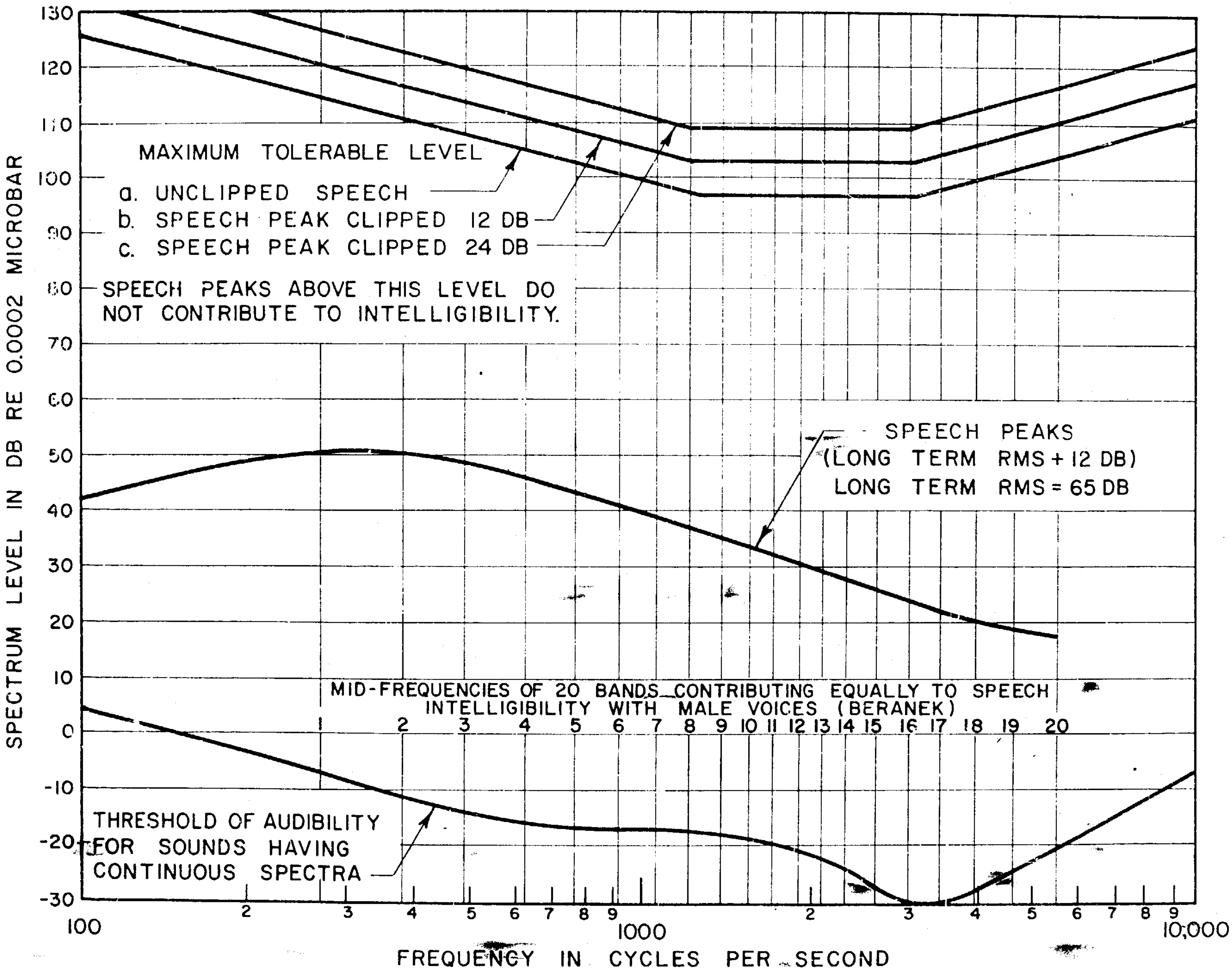

FIG. 4 WORK SHEET FOR AI - 20 BAND METHOD

the peak sound pressure level occurring in each word taken over a number of representative words or sentences as measured by a typical sound pressure level meter on C scale or on a properly calibrated rms volt meter. Accordingly, the speech peak curve on Fig. 4 can be moved up or down by the number of decibels the long-term rms of a given speech sample exceeds or falls short of 65 db. This procedure can be used only when there is present no appreciable amount of frequency distortion of the speech signal. When frequency distortion is a factor, the characteristics of speech spectrum must, of course, be directly measured.

Calculation of AI, Octave-Band Method

Each octave-band AI was calculated as follows: (a) the speech peak-to-noise ratio in each octave band was multiplied by the values given in column 3, Table 2; (b) the sum of the resulting numbers is the octave band AI.

The fractions used in Table 2 are derived from Beranek's formula for the 20-band Articulation Index. The weights in column 3, Table 2, reflect, in addition to the fraction .0017, the approximate number of the 20 bands of equal importance to speech intelligibility that appear in the respective octave bands. The approximate distribution of these 20 bands among the selected octave band is shown in Table 3. Coalescing the fraction .0017 with the weights for density of speech frequencies important to speech intelligibility gives the total weighting to be applied to the speech peak-to-noise ratios as found in the several octave bands (column 3, Table 2).

A work sheet to be used with the octave band method is presented in Fig. 5. The octave-band speech spectrum may be changed to meet the needs of a given system under test according to the procedures and restrictions set forth above with respect to work sheet for the 20-band method, Fig. 4.

Table 2. Procedure for Calculation of AI by Octave-Band Method.

1. Octave Band	2. Speech Peaks-to-Noise Ratio in Decibels*	3. Fractional Value of S/N Ratio for Intelligibility	4. Col 2 x Col 3
150-300 cps	________	.0013	________
300-600	________	.0042	________
600-1200	________	.0067	________
1200-2400	________	.0105	________
2400-4800	________	.0089	________
4800-9600	________	.0017	________
		AI = Σ	

* S/N ratios of 0 db or less are made 0; ratios of 30 db or greater are made 30.

Table 3. Number of Frequency Bands of Equal Importance to Intelligibility in Several Octave Bands.

Octave Band	Approximate Number of 20 Equally Important Frequency Bands
150-300 cps	1
300-600	2
600-1200	4
1200-2400	7
2400-4800	5
4800-9600	1

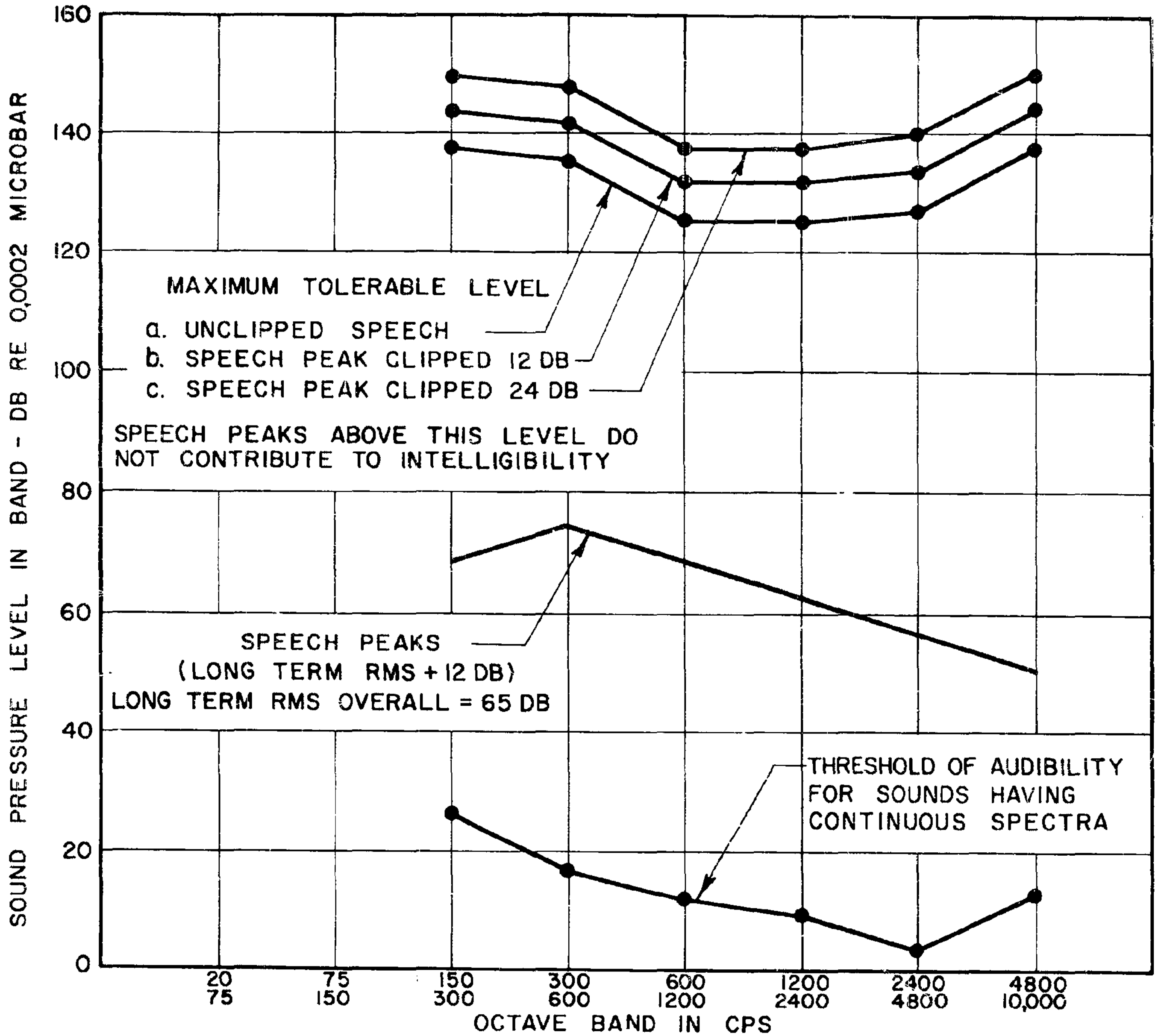

FIG. 5 WORK SHEET FOR AI - OCTAVE BAND METHOD

RESULTS AND DISCUSSION

The percentage of PB words correctly perceived by the listeners for the different test conditions is shown in Fig. 6. It is to be noted that, except perhaps for Noise B, the scores obtained for the two average noise levels of 80 and 105 db are about equal when presented at comparable speech-to-noise ratios. This, of course, agrees with many previous experiments of this sort. The ratios in Fig. 6 were determined from long-term rms measures of both the speech signal and the noise.

The difference between the two curves for Noise B could possibly be attributed to an increased spread of masking from the lower to the higher frequencies as the noise level was changed from 80 to 105 db. However, Noise B drops off in intensity above 600-1200 cps at the average rate of about 4-6 db per octave, and previous research (5) would indicate that the amount of masking due to a spreading effect would not equal local masking from the higher frequency components. It seems reasonable, therefore, to attribute the differences between the two functions for each of the Noises B, C, and D to experimental error and to strike an average for each pair of curves at the various signal-to-noise ratios.

20-Band AI's

AI's calculated according to the 20-band method for representative speech-to-noise ratios are plotted in Fig. 7 against the percent PB scores obtained with the same speech-to-noise ratios. It is to be noted that the points for the various noise conditions can be closely fitted by a single line. This would seem to demonstrate that the calculated AI's accurately predicted, relative to each other, the masking effects of the various noises. The absolute

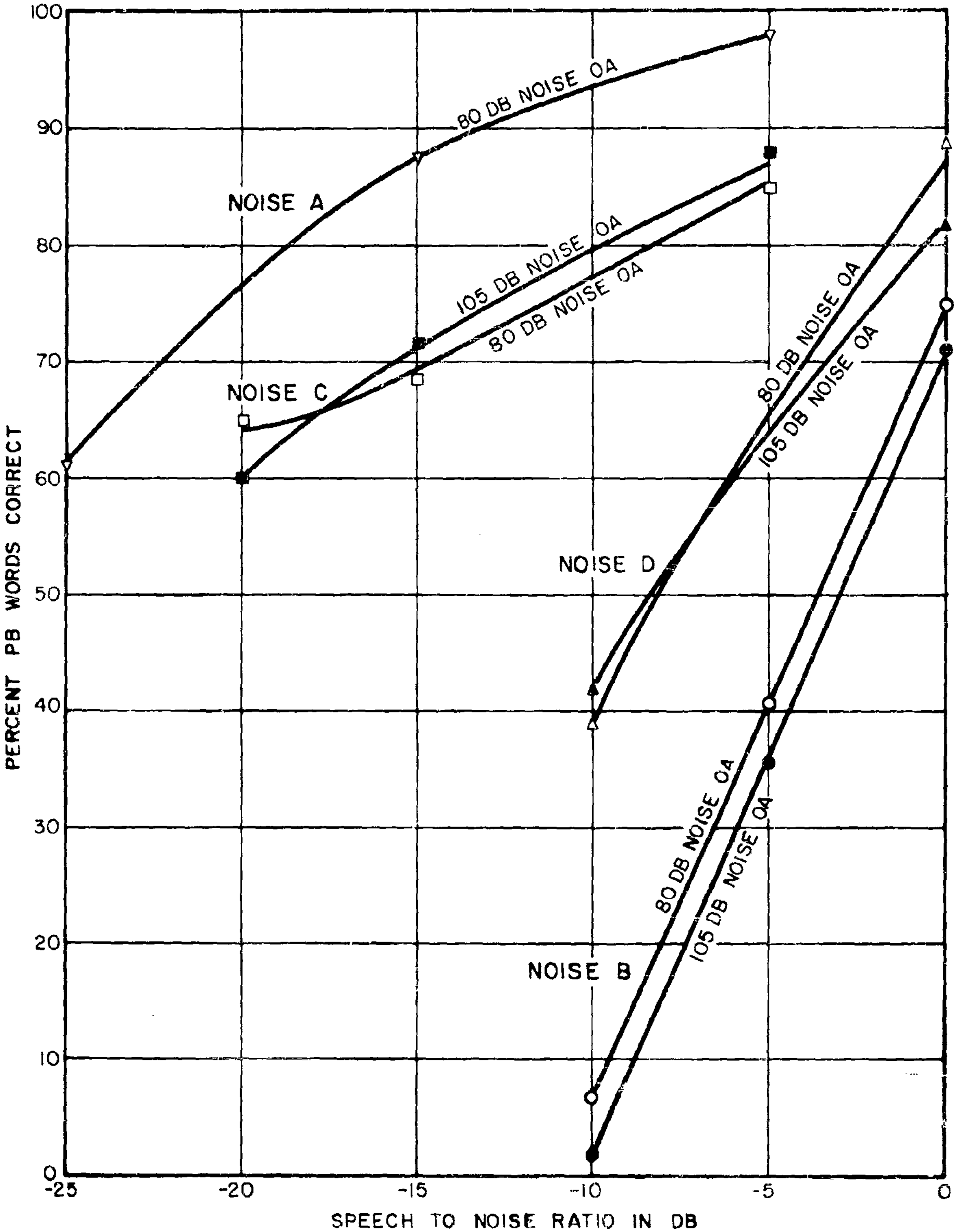

FIG. 6 PB WORD SCORES OBTAINED WITH DIFFERENT SIGNAL TO NOISE RATIOS AND NOISE LEVELS

value of the PB scores that were obtained are rather high compared to those usually reported for communication systems operating under comparable AI's as shown by the dashed curve in Fig. 7. The most probable reason is, as aforementioned, that the crew was so thoroughly experienced in taking PB word tests from the particular talker used.

Octave-Band AI's

The AI's obtained by the octave-band method are plotted in Fig. 8 against the intelligibility scores earned at various signal-to-noise ratios. Again, as with the 20-band method, the various plotted points are nicely fitted with a single line, indicating that the method adequately handles or predicts the masking effects of the different noises tested.

CONCLUSIONS

On the basis of the present experiment the following conclusions are drawn: AI's computed by either the 20-band or octave-band methods predict the results of the experiment herein reported. Presumably the octave-band method can be used instead of the longer, more complex 20-band method when: (a) the speech signal is relatively undistorted; and (b) when the noise is steady-state and does not have slopes radically different from those used in these tests.

The Articulation Index accurately predicts the masking effects of noises with continuous spectra the average slopes of which do not deviate more than +9 db or -9 db per octave. It is probable that the AI method will adequately handle noises having even somewhat steeper average slopes when the levels are somewhat lower, or at

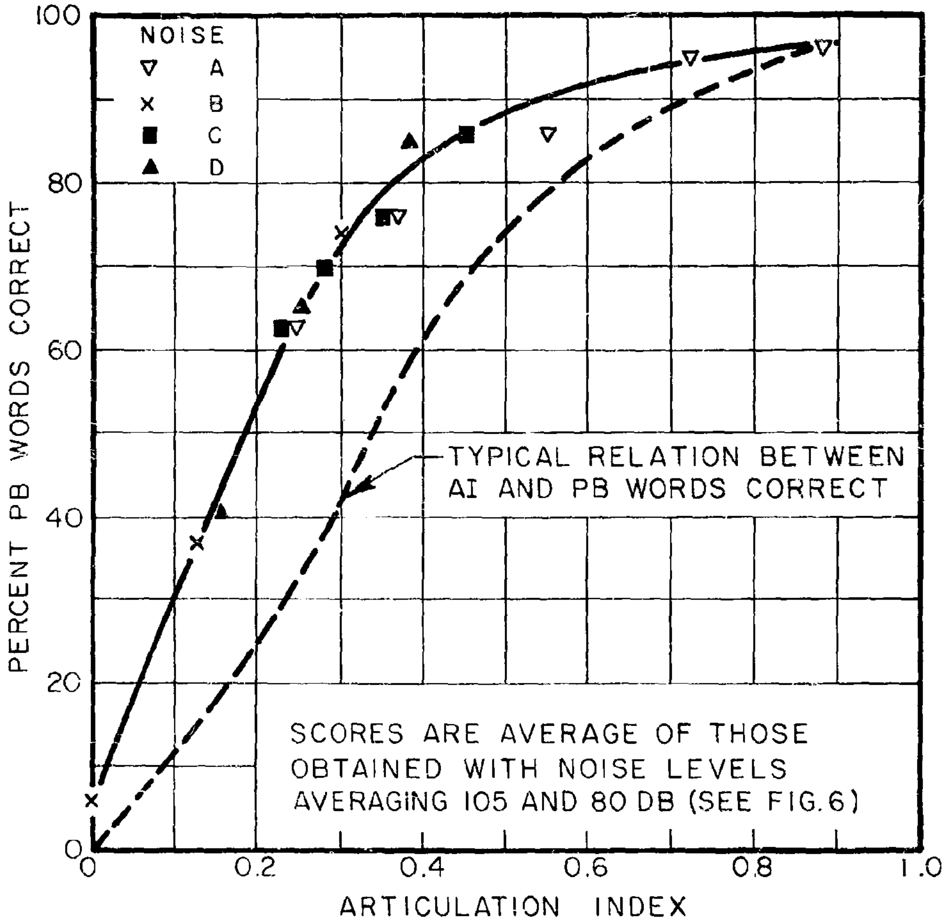

FIG.7 COMPARISON OF OBTAINED AND PREDICTED INTELLIGIBILITY TEST SCORES - 20 BAND METHOD

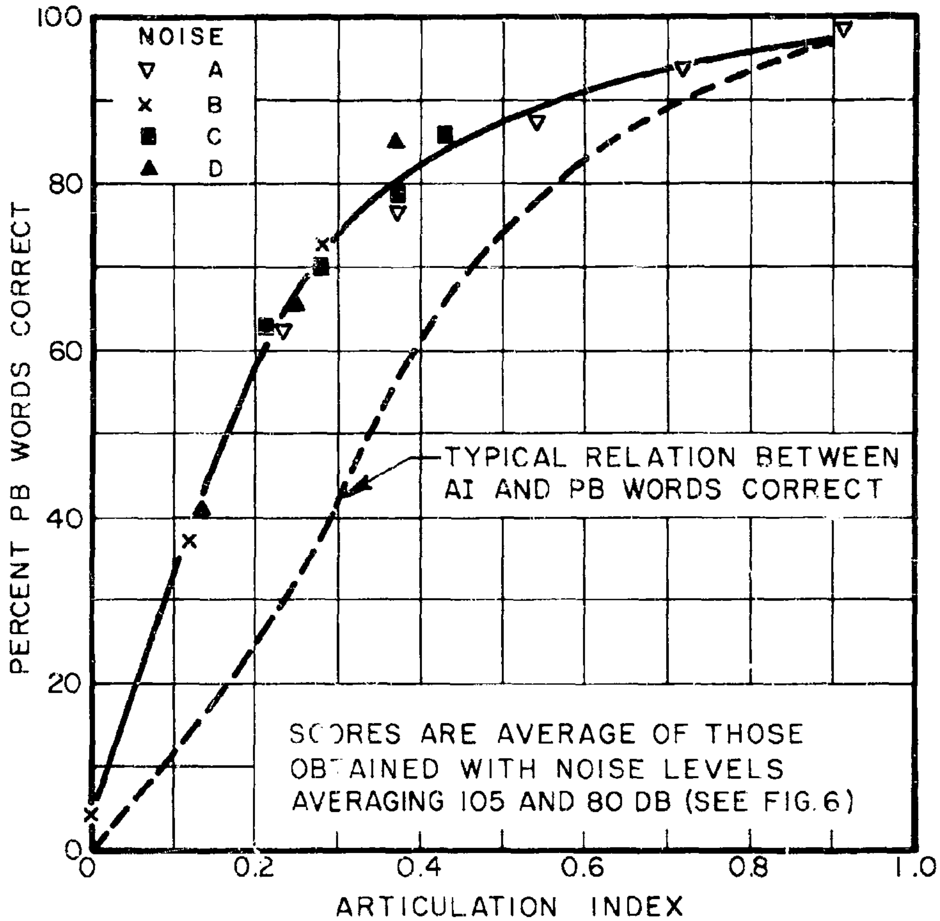

FIG.8 COMPARISON OF OBTAINED AND PREDICTED INTELLIGIBILITY TEST SCORES – OCTAVE BAND METHOD

least no greater, than those used in these tests; on the other hand these AI methods may not be adequate for noises having "peaked" spectra at intensity levels higher than those used in the study herein reported.

For the noises and absolute levels employed in these tests there would be little loss in the accuracy of the octave-band method by: (a) eliminating the band 150-300 cps and assigning its weight to the band 300-600 cps; and (b) dropping band 4800-9600 cps from consideration and increasing the weight of band 2400-4800 cps by a comparable amount.

REFERENCES

1. N. R. French and J. C. Steinberg, "Factors Governing the Intelligibility of Speech Sounds," J. Acoust. Soc. Am., 19, 90-119, 1949.

2. L. L. Beranek, "The Design of Speech Communication Systems," Proc. Instit. of Radio Engineers, 35, 880-890, 1947.

3. K. D. Kryter, "On Predicting the Intelligibility of Speech from Acoustical Measures," J. Speech and Hearing Disorders, 21, 208-217, 1956.

4. K. D. Kryter, "Variables Affecting Speech Communication in Noise," Proc. 2nd Internat'l. Congr. Acoust., New York: Am. Instit. Physics, 1957.

5. K. D. Kryter, "Human Engineering Principles for the Design of Speech Communication Systems," AFCRC TR 58-62, 1958, U. S. Dept. of Commerce, Office of Technical Services, Washington 25, D. C.

6. J. M. Pickett and K. D. Kryter, "Prediction of Speech Intelligibility in Noise," AFCRC TR 58-62, 1958, U. S. Dept. of Commerce, Office of Technical Services, Washington 25, D. C.

7. J. P. Egan and F. M. Wiener, "On the Intelligibility of Bands of Speech in Noise," J. Acoust. Soc. Am., 18, 435-441, 1946.

REFERENCES (Continued)

8. I. J. Hirsh, E. G. Reynolds, and M. Joseph, "Intelligibility of Different Speech Materials," J. Acoust. Soc. Am., 26, 530-538, 1954.

9. G. A. Miller, "The Masking of Speech," Psychol. Bull., 44, 105-129, 1947.

10. J. P. Egan and E. J. Thwing, Unpublished research, Hearing and Communication Laboratory, Indiana University, Bloomington, Indiana.

UNCLASSIFIED

UNCLASSIFIED